



Working Paper 14-21
Statistics and Econometrics Series (14)
July 2014

Departamento de Estadística
Universidad Carlos III de Madrid
Calle Madrid, 126
28903 Getafe (Spain)
Fax (34) 91 624-98-49

GOODNESS-OF-FIT TEST FOR RANDOMLY CENSORED DATA BASED ON MAXIMUM CORRELATION

Ewa Strzalkowska-Kominiak and Aurea Grané⁽¹⁾

Abstract

In this paper we study the goodness-of-fit test introduced by Fortiana and Grané (2003) and Grané (2012), in the context of randomly censored data. We construct a new test statistic under general right-censoring, i.e., with unknown censoring distribution, and prove its asymptotic properties. Additionally, we study a special case, when the censoring mechanism follows the well-known Koziol-Green model. We present an extensive simulation study on the empirical power of these two versions of the test statistic. We show the good performance of these statistics in detecting symmetrical alternatives and their advantages over the most famous Pearson-type test proposed by Akritas (1988). Finally, we apply our test to the head-and-neck cancer data.

Keywords: Goodness-of-fit; Kaplan-Meier estimator; Maximum correlation; Random censoring.

AMS classification: 62E20, 62G10, 62G20, 62N86.

(1) A. Grané, Statistics Department, Universidad Carlos III de Madrid, C/ Madrid 126, 28903 Getafe (Madrid), Spain, e-mail: aurea.grane@uc3m.es, corresponding author.
E. Strzalkowska-Kominiak, Statistics Department, Universidad Carlos III de Madrid, C/ Madrid 126, 28903 Getafe (Madrid), Spain, e-mail: ewa.strzalkowska@uc3m.es

This work has been partially supported by Spanish grants MTM2010-17323 (Spanish Ministry of Science and Innovation), MTM2011-22392, ECO2011-25706 (Spanish Ministry of Economy and Competitiveness).

Goodness-of-fit test for randomly censored data based on maximum correlation

Ewa Strzalkowska-Kominiak

Aurea Grané

*Statistics Department, Universidad Carlos III de Madrid.**

Abstract

In this paper we study the goodness-of-fit test introduced by Fortiana and Grané (2003) and Grané (2012), in the context of randomly censored data. We construct a new test statistic under general right-censoring, i.e., with unknown censoring distribution, and prove its asymptotic properties. Additionally, we study a special case, when the censoring mechanism follows the well-known Koziol-Green model. We present an extensive simulation study on the empirical power of these two versions of the test statistic. We show the good performance of these statistics in detecting symmetrical alternatives and their advantages over the most famous Pearson-type test proposed by Akritas (1988). Finally, we apply our test to the head-and-neck cancer data.

Keywords: Goodness-of-fit Kaplan-Meier estimator; Maximum correlation; Random censoring.

1 Introduction

In many medical studies one encounters data which are not fully observed and so censored from the right. Let Y_1, \dots, Y_n be the lifetimes of interest, coming from the continuous distribution function F and C_1, \dots, C_n the censoring times from the distribution function G . In the context of right-censored data, for every $i = 1, \dots, n$, we observe

$$X_i = \min(Y_i, C_i) \text{ and } \delta_i = 1_{\{Y_i \leq C_i\}},$$

where 1_A denotes the indicator function being equal to 1 if A is fulfilled and 0 otherwise. Even though the unknown distribution function F can be estimated by a well-known product-limit estimator, introduced by Kaplan and Meier (1958), very often we may have sufficient information to know the shape

*Statistics Department, Universidad Carlos III de Madrid, C/ Madrid 126, 28903 Getafe, Spain. E-mails: E. Strzalkowska-Kominiak, ewa.strzalkowska@uc3m.es, A. Grané, aurea.grane@uc3m.es (corresponding author).

of the distribution in question. Using such fully parametric models can lead to substantial gain in the efficiency of statistical procedures involving the distribution of the lifetimes F if only the parametric choice of that distribution is correct. This makes the goodness-of-fit test an important statistical tool, when dealing with (right-)censored data. It is clear, however, that the most famous tests for complete data, as Kolmogorov-Smirnov or Cramer-von Mises, are difficult to apply in the presence of censoring since the limit distribution depends on the censoring distribution G . The most recent overview of this kind of tests with randomly censored data is given by Balakrishnan et al (2014). Some more classical approaches are to be found in Koziol and Green (1976) and Akritas (1988). While the first one is based on the assumption that the distribution function G follows the so called Koziol-Green model and hence is more restrictive, the second one is a χ^2 test applied to general random censoring. For this reason, the Pearson-type goodness-of-fit test proposed by Akritas (1988) is so far the best option for randomly censored data with unknown censoring distribution. Nevertheless, it requires a partition of the observations to the cells and an adequate choice of number of classes since the power of the test may vary depending on the degrees of freedom. In this work we propose a new goodness-of-fit test based on maximum correlation coefficient which has a normal limiting distribution and hence it is straightforward to apply.

For this, we first introduce the maximum correlation in a more general set-up. Let Y_1 and Y_2 be two random variables with finite second order moments, joint cumulative distribution function (cdf) H and marginals F_1 and F_2 , respectively. Recall, that the Hoeffding representation of the correlation coefficient is given by

$$\rho(F_1, F_2) = \frac{1}{\sigma_1 \sigma_2} \int_{\mathbb{R}^2} (H(x, y) - F_1(x)F_2(y)) dx dy,$$

where σ_i denotes the standard deviation of Y_i . Furthermore, the maximum correlation of the pair of random variables (Y_1, Y_2) is defined as the correlation coefficient $\rho^+(F_1, F_2)$ corresponding to the bivariate cdf $H^+(x, y) = \min(F_1(x), F_2(y))$, the upper Fréchet bound of $H(x, y)$. The cdf $H^+(x, y)$ is a singular distribution, having support on the one-dimensional set $\{(x, y) \in \mathbb{R}^2 : F_1(x) = F_2(y)\}$, and the maximum correlation coefficient is given by

$$\rho^+(F_1, F_2) = \frac{1}{\sigma_1 \sigma_2} \left(\int_0^1 F_1^{-1}(p) F_2^{-1}(p) dp - \mu_1 \mu_2 \right), \quad (1)$$

where F_i^{-1} is the inverse of F_i and μ_i is the mean of Y_i . This maximum correlation, $\rho^+(F_1, F_2)$, is a measure of agreement between F_1 and F_2 , since $\rho^+ = 1$ if and only if $F_1 = F_2$ up to a scale and location change. In particular, Cuadras and Fortiana (1993) proposed the statistic based on $\rho^+(F, F_0)$ as a measure of goodness of fit of an iid sample Y_1, \dots, Y_n with cdf $F_1 = F$, to a given distribution F_0 . The goodness-of-fit test based on maximum correlation was further studied by Fortiana and Grané (2003) and Grané (2012).

As in the two latter publications, the present paper is devoted to testing uniformity, i.e. $F_0 = F_U$, a $[0, 1]$ uniform distribution. As shown by Fortiana and Grané (2003) the asymptotic approximation of

$\rho^+(F, F_U)$ is available, but convergence to its limiting law is rather slow. This led to definin

$$Q = \frac{\sigma}{\sqrt{1/12}} \rho^+(F, F_U) = 6 \int_0^1 x(2F(x) - 1)F(dx), \quad (2)$$

which equals one is $F = F_U$.

The goal of this paper is to study a test statistic based on Q when Y_1, \dots, Y_n may not be fully observed and so censored from the right by censoring times C_1, \dots, C_n . More precisely, we wish to test the hypothesis $H_0 : F = F_U$, where F_U is the cdf of a $[0, 1]$ uniform random variable, based on the sample $(X_i, \delta_i)_{i=1, \dots, n}$, where $X_i = \min(Y_i, C_i)$, with $X_i \in [0, 1]$. Nevertheless, our approach is not restricted to testing uniformity. We can also consider a more general null hypothesis F_0 since, under $H_0 : F = F_0$, we have that the transformed random variable $F_0(Y)$ follows a $[0, 1]$ uniform distribution. That is, $\tilde{Y} = F_0(Y) \sim F_U$ under the null hypothesis. Setting $\tilde{C} = F_0(C)$ and since $\{\tilde{Y}_i \leq \tilde{C}_i\} = \{Y_i \leq C_i\}$, leads us to testing uniformity based on the iid sample $(\tilde{X}_1, \delta_1), \dots, (\tilde{X}_n, \delta_n)$, where

$$\tilde{X}_i = \min(\tilde{Y}_i, \tilde{C}_i) \text{ and } \delta_i = 1_{\{\tilde{Y}_i \leq \tilde{C}_i\}}.$$

Hence, testing for uniformity is equivalent to testing for a fully specific continuous distribution. Even though it seems that we can extend the work of Fortiana and Grané (2003) setting $Q_n = 6 \int_0^1 x(2F_n(x) - 1)F_n(dx)$, where F_n denotes the Kaplan-Meier estimator for censored data, it is far from being true. In contrast to the empirical distribution under completely observed data, the Kaplan-Meier estimator is biased (see Stute (1994), for details). In Section 2 we show that such Plug-In estimator suffers from bias of this product-limit estimator and hence $E(Q_n) = 1$ under H_0 does not hold. To avoid this problem we propose to re-write Q in such a way, that it can be estimated by degenerated U-statistics. This leads to significant bias (and variance) reduction. In Section 3 we prove the asymptotic normality of the proposed estimator and in Section 4 we present our new goodness-of-fit test. In Section 5 we present an extensive simulation study. Finally, in Section 6 we adapt the test statistic to the case of composite null hypothesis and apply our test to the head-and-neck cancer data from Haghighi and Nikulin (2004).

2 Test statistic

In this section we propose our new statistic to test the goodness of fit for randomly censored data based on the modified maximum correlation coefficient. Recall that, under $H_0 : F = F_U$, the quantity

$$Q = \frac{\sigma}{\sqrt{1/12}} \rho^+(F, F_U) = 6 \int_0^1 x(2F(x) - 1)F(dx)$$

equals one. Hence in the following we prefer to work with

$$Q^1 = Q - 1 = 6 \int_0^1 x(2F(x) - 1)F(dx) - 1 \quad (3)$$

which equals zero if H_0 is true.

First, we define a Plug-In estimator of Q^1 by replacing F in (3) with the well-known Kaplan-Meier estimator. We obtain

$$Q_n^1 = 6 \int_0^1 x(2F_n(x) - 1)F_n(dx) - 1, \quad (4)$$

where F_n is defined as follows

$$F_n(x) = 1 - \prod_{X_i \leq x} \left[1 - \frac{\delta_i}{\sum_{k=1}^n 1_{\{X_k \geq X_i\}}} \right]. \quad (5)$$

It turns out, that under the null hypothesis and for finite samples the Plug-In estimator Q_n^1 suffers from significant bias and its convergence to the limiting distribution is very slow.

To solve this problem, we propose to estimate Q^1 with a U-statistic. For this, note that, if F is a continuous cdf and $\text{supp}(F) \subseteq [0, 1]$, then

$$2 \int_0^1 F(x)F(dx) = 1.$$

Hence

$$\begin{aligned} Q^1 &= \int_0^1 (6x(2F(x) - 1) - 2F(x))F(dx) = \int_0^1 [(6x - 2)F(x) - 6x(1 - F(x))]F(dx) \\ &= \int_0^1 \int_0^1 [(6x - 2)1_{\{y \leq x\}} - 6x1_{\{y > x\}}]F(dx)F(dy). \end{aligned} \quad (6)$$

Now we may replace the unknown quantities by their estimators. For this we introduce the jumps of the Kaplan-Meier estimator by setting

$$w_{in} = F_n(X_i) - F_n(X_i-),$$

where $F_n(x-)$ is the left-continuous version of $F_n(x)$.

Finally, the estimator of Q^1 is given by

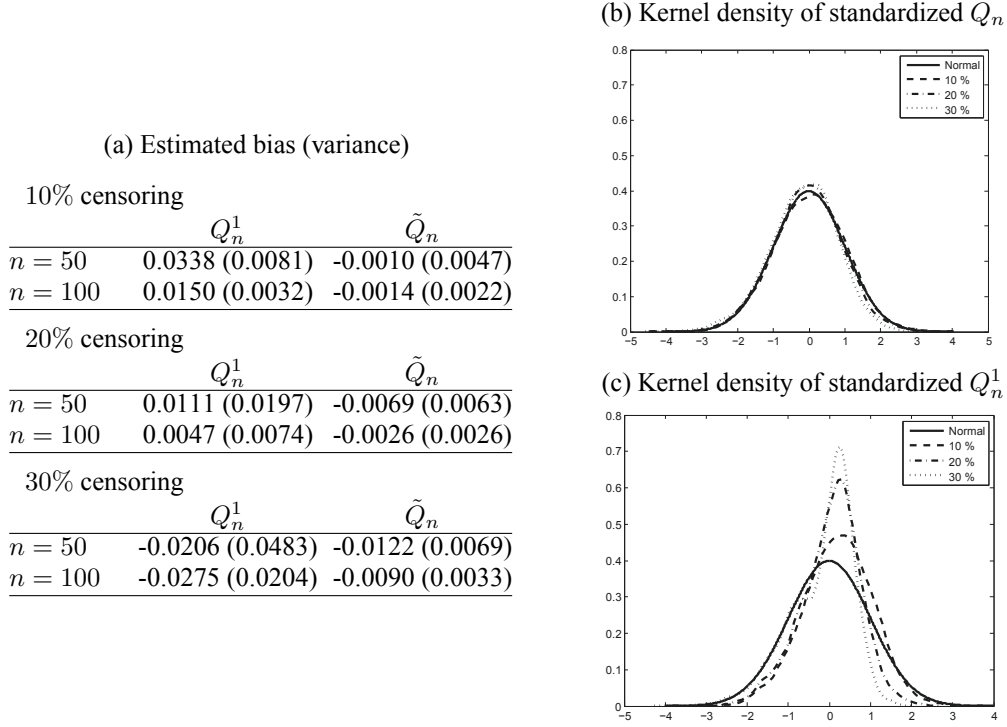
$$\tilde{Q}_n = \sum_{i=1}^n \sum_{j \neq i}^n w_{in} w_{jn} h(X_i, X_j), \quad (7)$$

where

$$h(x_1, x_2) = (6x_1 - 2)1_{\{x_2 \leq x_1\}} - 6x_1 1_{\{x_2 > x_1\}}.$$

To illustrate the advantages of using \tilde{Q}_n over the Plug-In estimator Q_n^1 , in panel (a) of Figure 1, we present the bias and variance of those estimators under the null hypothesis, that is, when the data comes from the $[0, 1]$ uniform distribution. Additionally, in panels (b)-(c) of Figure 1, we compare the kernel density estimators of the standardized versions of \tilde{Q}_n and Q_n^1 to that of the standard normal distribution. The standardization is done using the estimated asymptotic variances, discussed later on. Clearly, the U-statistic, \tilde{Q}_n , exhibits much smaller bias (and variance) and its standardized version fits nicely the standard normal distribution for all the considered censoring rates.

Figure 1: Comparison between \tilde{Q}_n and the Plug-In estimator Q_n^1 : Estimated bias (variance) based on 5000 trials and kKernel densities for $n = 200$.



3 Asymptotic properties

In this section we study the asymptotic properties of our test statistic \tilde{Q}_n . First we consider \tilde{Q}_n under general censoring mechanism, that is, without assuming any shape of the distribution function $G(x) = P(C \leq x)$, which is the cdf of the censoring times. Then we apply the results to the special case of Koziol-Green model. Recall that $F(x) = P(Y \leq x)$ is the cdf of the lifetimes of interest. We need following assumptions:

$$A1 : \int_0^1 \frac{F(du)}{1 - G(u)} < \infty$$

$$A2 : \int_0^1 |\varphi(u)| C^{1/2}(u) F(du) < \infty$$

where φ is a score function, $C(x) = \int_0^x \frac{G(dy)}{(1-G(y))^2(1-F(y))}$ and F is continuous with support in $[0, 1]$.

Theorem 1. Under A1 and A2, we have

$$\sqrt{n}(\tilde{Q}_n - Q^1) \rightarrow \mathcal{N}(0, \sigma^2),$$

where

$$\sigma^2 = \int_0^1 \frac{\varphi^2(x)}{1 - G(x)} F(dx) - \left[\int_0^1 \varphi(x) F(dx) \right]^2 - \int_0^1 \left[\int_x^1 \varphi(y) F(dy) \right]^2 \frac{(1 - F(x))G(dx)}{(1 - H(x))^2}$$

and

$$\varphi(x) = 12xF(x) - 6x - 2 - 12 \int_0^x yF(dy) + 6 \int_0^1 yF(dy).$$

Proof.

In view of (7), we can write \tilde{Q}_n in the following way

$$\tilde{Q}_n = \int_0^1 \int_0^1 \tilde{h}(x, y) F_n(dx) F_n(dy), \quad (8)$$

where

$$\tilde{h}(x, y) = (6x - 2)1_{\{y < x\}} - 6x1_{\{y > x\}}.$$

In the first step of the proof we write \tilde{Q}_n as a sum of four terms as follows

$$\tilde{Q}_n = \tilde{Q}_1 + \tilde{Q}_{2n} + \tilde{Q}_{3n} + \tilde{Q}_{4n},$$

where

$$\begin{aligned} \tilde{Q}_1 &= \int_0^1 \int_0^1 \tilde{h}(x, y) F(dx) F(dy) \\ \tilde{Q}_{2n} &= \int_0^1 \int_0^1 \tilde{h}(x, y) (F_n(dx) - F(dx)) F(dy) \\ \tilde{Q}_{3n} &= \int_0^1 \int_0^1 \tilde{h}(x, y) (F_n(dy) - F(dy)) F(dx) \\ \tilde{Q}_{4n} &= \int_0^1 \int_0^1 \tilde{h}(x, y) (F_n(dx) - F(dx)) (F_n(dy) - F(dy)). \end{aligned}$$

By (6) and since F is continuous, we have that $\tilde{Q}_1 = Q^1$. As to $\tilde{Q}_{2n} + \tilde{Q}_{3n}$, we obtain

$$\tilde{Q}_{2n} + \tilde{Q}_{3n} = \int_0^1 \varphi(x) (F_n(dx) - F(dx)),$$

where

$$\begin{aligned} \varphi(x) &= \int_0^1 \tilde{h}(x, y) F(dx) + \int_0^1 \tilde{h}(x, y) F(dy) \\ &= 12xF(x) - 6x - 2 - 12 \int_0^x yF(dy) + 6 \int_0^1 yF(dy). \end{aligned}$$

It remains to show that $\tilde{Q}_{4n} = o_{\mathbb{P}}(n^{1/2})$. For this, set $\tau_{\tilde{H}} = \inf\{t : \tilde{H}(t) = 1\}$, where $\tilde{H}(t) = \mathbb{P}(X \leq t)$ is the cdf of the observed sample. Since $\text{supp}(F) \in [0, 1]$ and G fulfill assumption A1, we have that $\tau_{\tilde{H}} = 1$. Moreover, by definition of $\tilde{h}(x, y)$, we can show that

$$\tilde{Q}_{4n} = -12 \int_0^1 x(F_n(x) - F(x))(F_n(dy) - F(dy)) - 2(F_n(1) - F(1))^2 =: \tilde{Q}_{4n}^a + \tilde{Q}_{4n}^b.$$

Now, we may consider the two terms, \tilde{Q}_{4n}^a and \tilde{Q}_{4n}^b , separately. According to Theorem 2 (7) in Ying (1989) and under A1, the process $\sqrt{n}(F_n - F)$ converges weakly to a Brownian process. See, also equation (11) in Wellner (2007). More precisely,

$$\sqrt{n}(F_n - F) \rightarrow (1 - F)\mathbb{B}(C), \text{ in } D[0, \tau_{\tilde{H}}],$$

where $D[0, \tau_{\tilde{H}}]$ denotes the Skorohod space. Furthermore, since F is continuous and D^0 is a set of uniformly bounded functions, we have that $\sqrt{n}(F_n - F) \in D^0$ with probability exceeding $1 - \varepsilon$ for every $\varepsilon > 0$. Additionally, $x \in [0, 1]$ and $\sup_{x \in [0, \tau_{\tilde{H}}]} |F_n(x) - F(x)| \rightarrow 0$ almost surely. Hence, using Theorem 2.1 in Rao (1962) with $g(x) = \sqrt{n}(F_n(x) - F(x))x$, we obtain

$$\sqrt{n}\tilde{Q}_{4n}^a = -12 \int_0^1 g(x)(F_n(dy) - F(dy)) = o_{\mathbb{P}}(1).$$

Additionally, under A1, $\sqrt{n}\tilde{Q}_{4n}^b = o_{\mathbb{P}}(1)$.

Finally, we obtain the following representation

$$\tilde{Q}_n = Q^1 + \int_0^1 \varphi(x)(F_n(dx) - F(dx)) + o_{\mathbb{P}}(n^{1/2}).$$

The asymptotic normality is now a direct consequence of Stute (1995). More precisely, under A1 and A2, we obtain

$$\sqrt{n} \int_0^1 \varphi(x)(F_n(dx) - F(dx)) \rightarrow \mathcal{N}(0, \sigma^2).$$

This completes the proof. \(\square\)

Consequently, we have

Corollary 1. *Under H_0 , A1 and A2, we have*

$$\sqrt{n}\tilde{Q}_n \rightarrow \mathcal{N}(0, \sigma^2).$$

It is to see, that the variance under H_0 would not simplify since it does depend on the distribution function of the censoring times G , which is unknown. Nevertheless, under the Koziol-Green model, we have an explicit expression for σ^2 . First, recall that G follows a Koziol-Green model if

$$1 - G(x) = (1 - F(x))^\beta,$$

where β is an unknown parameter. We can see, however, that

$$p := P(Y > C) = \frac{\beta}{\beta + 1} \text{ and } 1 - p = \int (1 - G(x))F(dx).$$

Hence β can be easily estimated using Kaplan-Meier estimators for F and G . Finally, it is easy to check that the assumptions A1 and A2 are fulfilled under the Koziol-Green model with $\beta \in (0, 1)$, that is, if the censoring is not heavier than 50%, which is a very reasonable assumption. So, as a consequence of Corollary 1, we get the following result.

Corollary 2. *Under Koziol-Green model with $\beta \in (0, 1)$ we have that, under H_0 ,*

$$\sqrt{n}\tilde{Q}_n \rightarrow \mathcal{N}(0, \sigma_{KG}^2),$$

where

$$\sigma_{KG}^2 = \frac{-\beta^4 + 4\beta^3 - 17\beta^2 + 38\beta - 24}{(\beta - 1)(\beta - 2)(\beta - 3)(\beta - 4)(\beta - 5)}.$$

4 Goodness-of-fi test

Once the test statistic is proposed and its limiting distribution is established, we are in the position to define the goodness-of-fi test. For this we estimate the asymptotic variance σ^2 using the Plug-In principle. That is, we replace the unknown quantities with its estimators. First, we define the distribution function of the observed times $\tilde{H}(x) = P(X \leq x)$ and set $\tilde{H}_n(x) = \frac{1}{n} \sum_{i=1}^n 1_{\{X_i \leq x\}}$ as its empirical counterpart. Moreover, let

$$H^0(x) = \mathbb{P}(X \leq x, \delta = 0) = \int_0^x (1 - F(u))G(du)$$

and

$$H^1(x) = \mathbb{P}(X \leq x, \delta = 1) = \int_0^x (1 - G(u))F(du)$$

be the subdistributions related with the observed censored and uncensored lifetimes. Their estimators are defined as follows

$$H_n^0(x) = \frac{1}{n} \sum_{i=1}^n 1_{\{X_i \leq x\}}(1 - \delta_i)$$

and

$$H_n^1(x) = \frac{1}{n} \sum_{i=1}^n 1_{\{X_i \leq x\}}\delta_i.$$

Hence

$$\begin{aligned} \sigma_n^2 &= \frac{1}{n} \sum_{i=1}^n \frac{\varphi_n^2(X_i)}{(1 - G_n(X_i-))^2} \delta_i - \left[\frac{1}{n} \sum_{i=1}^n \frac{\varphi_n(X_i)}{1 - G_n(X_i-)} \delta_i \right]^2 \\ &\quad - \frac{1}{n} \sum_{i=1}^n \frac{1 - \delta_i}{(1 - H_n(X_i-))^2} \left[\frac{1}{n} \sum_{j=1}^n \frac{\varphi_n(X_j)}{1 - G_n(X_j-)} \delta_j 1_{\{X_j \geq X_i\}} \right]^2, \end{aligned}$$

where

$$\varphi_n(x) = 12xF_n(x) - 6x - 2 - 12\frac{1}{n} \sum_{i=1}^n \frac{X_i\delta_i}{1 - G_n(X_i-)} 1_{\{X_i \leq x\}} + 6\frac{1}{n} \sum_{i=1}^n \frac{X_i\delta_i}{1 - G_n(X_i-)}.$$

and G_n is a Kaplan-Meier estimator given by

$$1 - G_n(x) = \prod_{X_i \leq x} \left[1 - \frac{1 - \delta_i}{\sum_{k=1}^n 1_{\{X_k \geq X_i\}}} \right].$$

Finally, we set

$$T_n := \frac{\sqrt{n}\tilde{Q}_n}{\sqrt{\sigma_n^2}}. \quad (9)$$

We reject H_0 at level α if

$$T_n \leq \Phi^{-1}(\alpha/2) \text{ or } T_n \geq \Phi^{-1}(1 - \alpha/2),$$

where Φ^{-1} is the inverse of the standard normal cdf.

Additionally, under the Koziol-Green model and in view of Corollary 2, we defin

$$T_n^{KG} := \frac{\sqrt{n}\tilde{Q}_n}{\sqrt{\hat{\sigma}_{KG}^2}}, \quad (10)$$

where

$$\hat{\sigma}_{KG}^2 = \frac{-\hat{\beta}^4 + 4\hat{\beta}^3 - 17\hat{\beta}^2 + 38\hat{\beta} - 24}{(\hat{\beta} - 1)(\hat{\beta} - 2)(\hat{\beta} - 3)(\hat{\beta} - 4)(\hat{\beta} - 5)}$$

and

$$\hat{\beta} = \left(\int (1 - G_n(x)) F_n(dx) \right)^{-1} - 1.$$

As before we reject H_0 at level α if

$$T_n^{KG} \leq \Phi^{-1}(\alpha/2) \text{ or } T_n^{KG} \geq \Phi^{-1}(1 - \alpha/2).$$

5 Simulation study

Here we conduct an extensive simulation study to show the behavior of our test. In the following subsection we consider only the null hypothesis, while in Subsection 5.2 we include the power study under different alternatives. In both subsections we compare our method with the Pearson-type goodness-of-fit test proposed by Akritas (1988). Following the notation of the previous section, we denote by T_n and T_n^{KG} our test statistics for the general censoring and under the Koziol-Green model, respectively. See, equations (9) and (10) for details. Moreover, we denote by $A_{(nc)}$ the χ^2 test proposed by Akritas (1988), where nc denotes the number of cells.

5.1 Null hypothesis

In this section we present the results of the proposed methods under the null hypothesis and at 5% significance level. As mentioned before, we consider our tests T_n and T_n^{KG} , together with the test presented by Akritas (1988). Following the latter work, we consider $nc = 2$ and $nc = 5$ and denote these tests by $A_{(2)}$ and $A_{(5)}$, respectively. The results are based on 5000 trials. As it is to see in the Table 1, our test T_n and the one from Akritas hold very well the significance level. As to the one based on Koziol-Green model, it holds the 5% level when censoring is low but for more than 20% of missing data, the variance σ_{KG}^2 does not captures the variability of our \tilde{Q}_n correctly and hence the significance level is slightly overestimated for heavy censoring.

Table 1: Null Hypothesis

	10% censoring				20% censoring				30% censoring			
	T_n	$A_{(5)}$	$A_{(2)}$	T_n^{KG}	T_n	$A_{(5)}$	$A_{(2)}$	T_n^{KG}	T_n	$A_{(5)}$	$A_{(2)}$	T_n^{KG}
$n = 50$	0.0508	0.0552	0.0538	0.0578	0.0450	0.0560	0.0558	0.0530	0.0496	0.0596	0.0560	0.0648
$n = 100$	0.0480	0.0582	0.0548	0.0532	0.0478	0.0530	0.0492	0.0560	0.0502	0.0568	0.0528	0.0708
$n = 200$	0.0468	0.0524	0.0574	0.0498	0.0522	0.0508	0.0508	0.0604	0.0494	0.0522	0.0468	0.0666

5.2 Power study

To study the power of our test we consider two different models:

Model 1: To test the uniformity ($H_0 : F = F_U$) we choose three parametric families of alternative probability distributions with support on $[0, 1]$.

(a) Lehmann alternatives,

$$F_\theta(x) = x^\theta, 0 \leq x \leq 1, \theta > 0;$$

where for $\theta = 1$ we have $F_\theta = F_U$.

(b) compressed uniform alternatives,

$$F_\theta(x) = \frac{x - \theta}{1 - 2\theta}, \theta \leq x \leq 1 - \theta,$$

where $0 \leq \theta \leq 1/2$; and for $\theta = 0$ we have $F_\theta = F_U$.

(c) centered distributions having a U-shaped density for $\theta \in (0, 1)$, or wedge-shaped density for $\theta > 1$

$$F_\theta(x) = \begin{cases} \frac{1}{2}(2x)^\theta, & 0 \leq x \leq 1/2 \\ 1 - \frac{1}{2}(2(1-x))^\theta, & 1/2 \leq x \leq 1 \end{cases}$$

where for $\theta = 1$ we have $F_\theta = F_U$.

Model 2: An exponentiality test (with parameter $\lambda = 1$), where the alternatives are Weibull distributions with parameters 1 and θ . More precisely, $F_\theta(x) = 1 - e^{-x^\theta}$, where $\theta = 1$ gives us the exponential distribution of the null hypothesis.

Additionally, the censoring variable, C , is generated under the Koziol-Green model. That is, $1 - G(x) = (1 - F(x))^\beta$, where $\beta = \frac{p}{1-p}$ and $p = \mathbb{P}(X > C)$ is the censoring level.

In the following figure and tables we present the power study at a 5% significance level. Panels (a1)-(c3) of Figure 2 contain the power of the test for Model 1 and panels (d1)-(d3) of Figure 2 contains the power under Model 2, for different sample sizes ($n = 50, 100, 200$) and one censoring level of 20%. All those figure are based on 2000 trials. Moreover, Tables 2 are based on 5000 trials and show the power under alternatives for two different sample sizes ($n = 100, 200$), censoring levels $p = 0.1, 0.2, 0.3$ and different values of the parameters θ . In particular, for Model 2, we choose

$\theta = 1 + Hn^{-0.5}$ and $H \in \{-4, -2, 2, 4\}$. Both, tables and figures include a comparison to the Pearson-type test proposed by Akritas (1988). As before, we use the number of cells (nc) equal to 2 and 5.

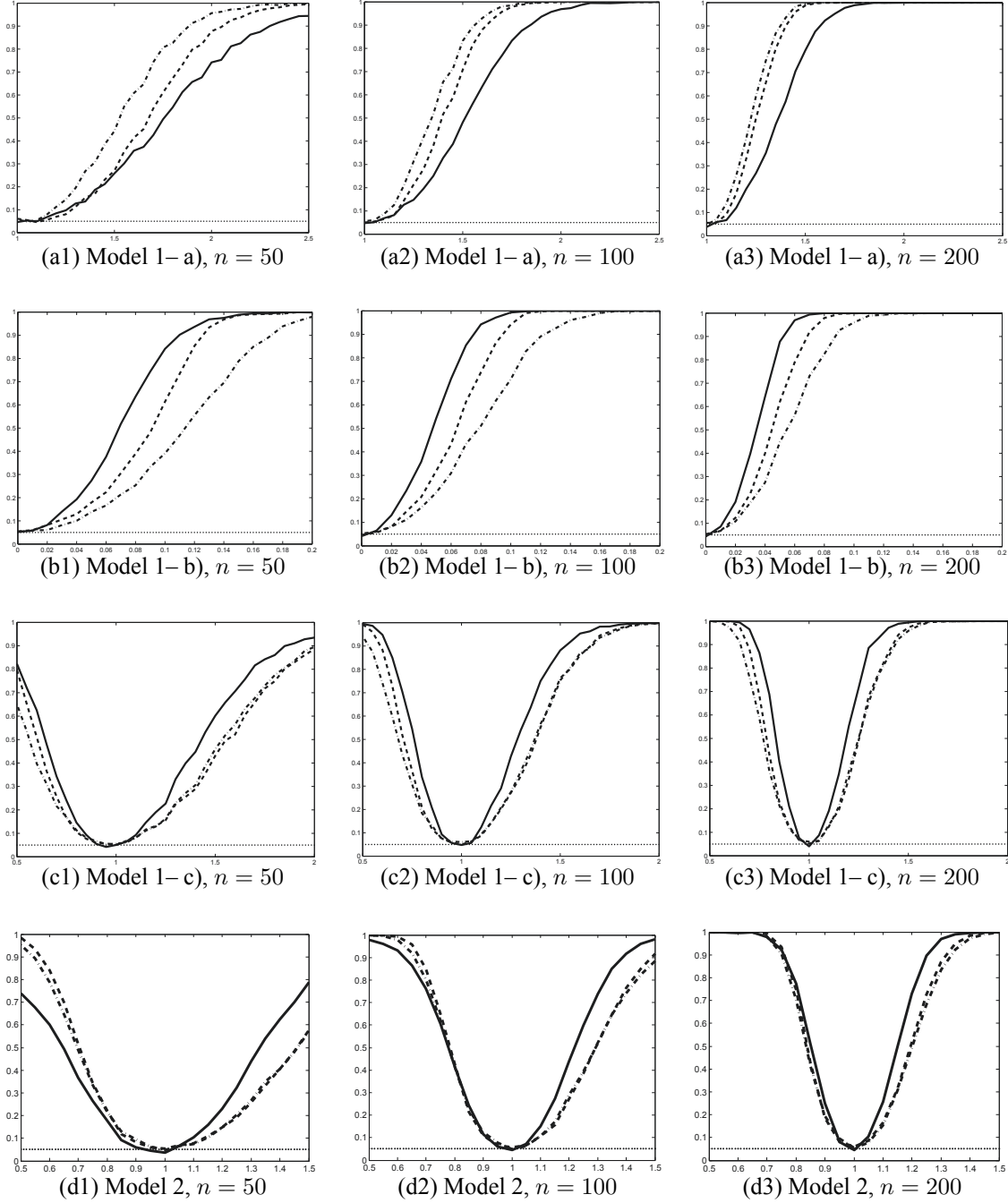


Figure 2: Power study for Model 1 (a1–c3) and Model 2 (d1–d3) for three different sample sizes and censoring rate $p = 0.2$. T_n (solid line), $A_{(5)}$ (dashed line) and $A_{(2)}$ (dash-dotted line).

Concerning the uniformity test (Model 1), it is clear that for alternatives b) and c) our test outperforms

the one proposed by Akritas. Additionally, our test does not depend on the number of cells and the choice of cell boundaries. The influence of the number of cells is made obvious in panels (a1)-(c3) of Figure 2. While $A_{(2)}$ gives better results than $A_{(5)}$ in alternative a), the opposite can be observed in alternatives b) and c). Regarding the exponentiality test (Model 2), we get better results than the competitive test of Akritas (1988) when the alternative is Weibull with parameter $\theta > 1$. For $\theta < 1$, our test reaches the high power of the Pearson-type test for big sample sizes. Notice, however, that in Model 2 and for all parameters θ , the test statistic under the Koziol-Green model, T_n^{KG} , gives very good results independently on the sample size.

Table 2: Power study for Model 1 and Model 2.

Model 1, Alternative a)												
$n = 100$	$p = 0.1$				$p = 0.2$				$p = 0.3$			
	T_n	$A_{(5)}$	$A_{(2)}$	T_n^{KG}	T_n	$A_{(5)}$	$A_{(2)}$	T_n^{KG}	T_n	$A_{(5)}$	$A_{(2)}$	T_n^{KG}
$\theta = 1.5$	0.5277	0.7107	0.8040	0.5547	0.4760	0.7220	0.8190	0.5167	0.3957	0.7093	0.8260	0.4570
$\theta = 2$	0.9847	0.9997	0.9997	0.9873	0.9997	1	1	1	0.9070	0.9980	0.9995	0.9520
$\theta = 2.5$	1	1	1	1	0.9997	1	1	1	0.9237	0.9990	0.9997	0.9607
$n = 200$	$p = 0.1$				$p = 0.2$				$p = 0.3$			
	T_n	$A_{(5)}$	$A_{(2)}$	T_n^{KG}	T_n	$A_{(5)}$	$A_{(2)}$	T_n^{KG}	T_n	$A_{(5)}$	$A_{(2)}$	T_n^{KG}
$\theta = 1.5$	0.8410	0.9823	0.9887	0.8540	0.8020	0.9830	0.9893	0.8200	0.7057	0.9833	0.9910	0.7517
$\theta = 2$	0.9997	1	1	0.9997	1	1	1	1	0.9950	1	1	0.9990
$\theta = 2.5$	1	1	1	1	1	1	1	1	0.9983	1	1	0.9997
Model 1, Alternative b)												
$n = 100$	$p = 0.1$				$p = 0.2$				$p = 0.3$			
	T_n	$A_{(5)}$	$A_{(2)}$	T_n^{KG}	T_n	$A_{(5)}$	$A_{(2)}$	T_n^{KG}	T_n	$A_{(5)}$	$A_{(2)}$	T_n^{KG}
$\theta = 0.05$	0.5987	0.3897	0.2940	0.5473	0.5370	0.3027	0.2297	0.5440	0.4563	0.2530	0.1803	0.5037
$\theta = 0.1$	1	0.9837	0.8510	0.9987	0.9957	0.9413	0.7163	0.9950	0.9410	0.9000	0.5910	0.9820
$\theta = 0.15$	1	1	0.9963	1	1	1	0.9797	1	0.9940	1	0.9500	1
$n = 200$	$p = 0.1$				$p = 0.2$				$p = 0.3$			
	T_n	$A_{(5)}$	$A_{(2)}$	T_n^{KG}	T_n	$A_{(5)}$	$A_{(2)}$	T_n^{KG}	T_n	$A_{(5)}$	$A_{(2)}$	T_n^{KG}
$\theta = 0.05$	0.9133	0.7517	0.5820	0.8830	0.8560	0.6003	0.4250	0.8313	0.7683	0.4870	0.3287	0.7973
$\theta = 0.1$	1	1	0.9933	1	1	1	0.9570	1	0.9977	0.9987	0.8977	1
$\theta = 0.15$	1	1	1	1	1	1	1	1	1	1	0.9993	1
Model 1, Alternative c)												
$n = 100$	$p = 0.1$				$p = 0.2$				$p = 0.3$			
	T_n	$A_{(5)}$	$A_{(2)}$	T_n^{KG}	T_n	$A_{(5)}$	$A_{(2)}$	T_n^{KG}	T_n	$A_{(5)}$	$A_{(2)}$	T_n^{KG}
$\theta = 0.75$	0.6587	0.3623	0.3217	0.6803	0.5337	0.3410	0.2877	0.6117	0.3490	0.3167	0.2823	0.5133
$\theta = 1.25$	0.4903	0.3090	0.3230	0.5347	0.4230	0.2667	0.2757	0.4957	0.3740	0.2323	0.2423	0.4720
$\theta = 1.5$	0.9513	0.8050	0.7987	0.9643	0.8787	0.7287	0.7353	0.9333	0.7907	0.6720	0.6800	0.9110
$n = 200$	$p = 0.1$				$p = 0.2$				$p = 0.3$			
	T_n	$A_{(5)}$	$A_{(2)}$	T_n^{KG}	T_n	$A_{(5)}$	$A_{(2)}$	T_n^{KG}	T_n	$A_{(5)}$	$A_{(2)}$	T_n^{KG}
$\theta = 0.75$	0.9337	0.7217	0.6167	0.9423	0.8933	0.6783	0.5860	0.9167	0.7307	0.6363	0.5423	0.8223
$\theta = 1.25$	0.7920	0.5513	0.5537	0.8103	0.7280	0.4823	0.4993	0.7773	0.6243	0.4277	0.4437	0.7303
$\theta = 1.5$	0.9993	0.9817	0.9750	0.9993	0.9980	0.9740	0.9623	0.9993	0.9613	0.9433	0.9233	0.9937
Model 2. Power study for $\theta = 1 + Hn^{-0.5}$												
$n = 100$	$p = 0.1$				$p = 0.2$				$p = 0.3$			
	T_n	$A_{(5)}$	$A_{(2)}$	T_n^{KG}	T_n	$A_{(5)}$	$A_{(2)}$	T_n^{KG}	T_n	$A_{(5)}$	$A_{(2)}$	T_n^{KG}
$H = -4$	0.9720	0.9960	0.9872	0.9914	0.9346	0.9928	0.9754	0.9844	0.7752	0.9822	0.9564	0.9328
$H = -2$	0.5580	0.4828	0.4746	0.6094	0.4236	0.4118	0.3966	0.5528	0.2354	0.3684	0.3510	0.4364
$H = 2$	0.4784	0.3112	0.3118	0.4954	0.4278	0.2608	0.2736	0.4730	0.3910	0.2194	0.2302	0.4578
$H = 4$	0.9594	0.8464	0.8296	0.9598	0.9160	0.7564	0.7430	0.9412	0.8204	0.6572	0.6478	0.9056
$n = 200$	$p = 0.1$				$p = 0.2$				$p = 0.3$			
	T_n	$A_{(5)}$	$A_{(2)}$	T_n^{KG}	T_n	$A_{(5)}$	$A_{(2)}$	T_n^{KG}	T_n	$A_{(5)}$	$A_{(2)}$	T_n^{KG}
$H = -4$	0.9876	0.9938	0.9856	0.9940	0.9704	0.9820	0.9604	0.9864	0.8664	0.9682	0.9306	0.9490
$H = -2$	0.5712	0.4390	0.4406	0.5976	0.4636	0.3870	0.3808	0.5312	0.2892	0.3370	0.3338	0.4268
$H = 2$	0.5108	0.3312	0.3354	0.5284	0.4562	0.2768	0.2818	0.4962	0.3998	0.2302	0.2410	0.4736
$H = 4$	0.9686	0.8946	0.8610	0.9676	0.9446	0.8216	0.7830	0.9550	0.8764	0.7330	0.7032	0.9316

6 Further extensions and Application

6.1 Composite Null Hypothesis

So far, our test T_n has been designed to test a fully specific null hypothesis. It does strongly base on the fact that the transformed lifetime $F_0(X)$ is $[0, 1]$ uniformly distributed under $H_0 : F = F_0$. In this section we consider a more general case, when the distribution function to be tested depends on an unknown parameter λ . Let now consider the following null hypothesis

$$H_0 : F \in \{F_\lambda : \lambda \in \mathbb{R}^d\}.$$

In this case first we need to estimate the parameter λ using, e.g., a maximum-likelihood estimator $\hat{\lambda}$. Clearly, if F_λ is twice differentiable in λ and the estimator $\hat{\lambda}$ is \sqrt{n} consistent, by the Taylor expansion we have that $F_{\hat{\lambda}}(X) = U + O_P(n^{-1/2})$, where $U = F_\lambda(X) \sim U[0, 1]$ under the null hypothesis H_0 . The test statistic \tilde{Q}_n should still admit a normal limit but the error term enters the variance of our test statistic and hence the asymptotic variance given in Theorem 1 is no longer valid. Even though the theoretical properties of our test in the case of such a composite hypothesis are beyond the scope of this paper, to test this kind of hypothesis we propose a modified test with jackknife estimator of the variance, which does take into account the estimation of parameters and works very well in practice. We proceed as follows:

1. Based on the sample X_1, \dots, X_n , find the maximum-likelihood estimator (MLE) $\hat{\lambda}$.
2. Define the pseudo-values $\tilde{X}_i = F_{\hat{\lambda}}(X_i)$ for $i = 1, \dots, n$.
3. Based on the sample $\tilde{X}_1, \dots, \tilde{X}_n$, compute the test statistic \tilde{Q}_n defined in (7).
4. Compute the jackknife estimator of the variance following the steps
 - For every $i = 1, \dots, n$, choose the subsample $X_1, \dots, X_{i-1}, X_{i+1}, \dots, X_n$ and compute the MLE $\hat{\lambda}^{(-i)}$.
 - Define the pseudo-values $\tilde{X}_j = F_{\hat{\lambda}^{(-i)}}(X_j)$ for $j = 1, \dots, i-1, i+1, \dots, n$.
 - Based on the sample $\tilde{X}_1, \dots, \tilde{X}_{i-1}, \tilde{X}_{i+1}, \dots, \tilde{X}_n$, compute the test statistic $\tilde{Q}_n^{(-i)}$.
 - Set

$$nV_n(\tilde{Q}_n) = (n-1) \sum_{i=1}^n (\tilde{Q}_n^{(-i)} - \bar{Q}_n)^2,$$

$$\text{where } \bar{Q}_n = \frac{1}{n} \sum_{i=1}^n \tilde{Q}_n^{(-i)}.$$

5. Define the test statistic

$$J_n := \frac{\sqrt{n}\tilde{Q}_n}{\sqrt{nV_n(\tilde{Q}_n)}}.$$

6. Reject H_0 , if

$$J_n \leq \Phi^{-1}(\alpha/2) \text{ or } J_n \geq \Phi^{-1}(1 - \alpha/2).$$

To check the behavior of this new jackknife-test, J_n , we study the hypothesis $H_0 : F \in \{\exp(\lambda) : \lambda \in (0, \infty)\}$, where the alternatives come from the Weibull distribution. Our simulated sample comes from $\exp(\lambda = 1)$ but λ is estimated using Maximum-Likelihood method. In Figure 3 we compare our test T_n from Section 4 with the test based on J_n .

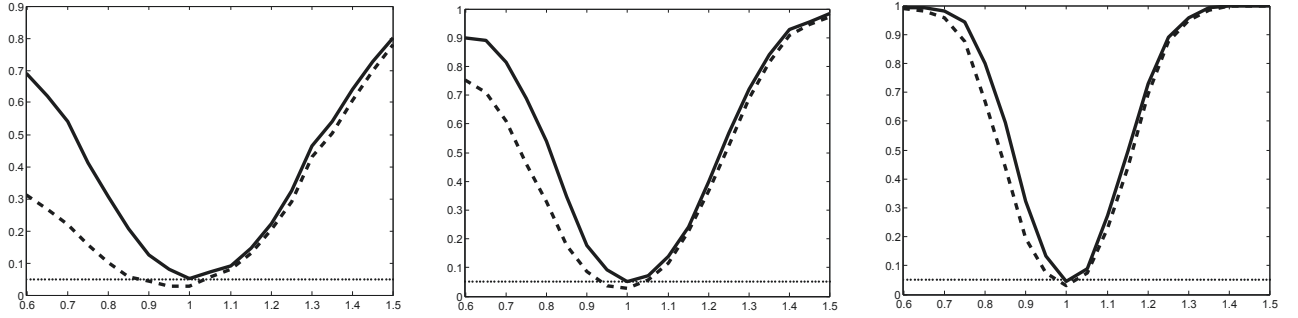


Figure 3: Power study: J_n (solid line) and T_n (dashed line), where $n = 50$ (left), $n = 100$ (middle) and $n = 200$ (right), $p = 0.2$

6.2 Real Data Example

We illustrate the use of our test on the head-and-neck cancer data from Haghighi and Nikulin (2004). These authors fitted the Power Generalized Weibull distribution $F(x, \sigma, v, \gamma)$ to the data. We use here the truncated data set with Survival Times (in months) smaller than 20. It was motivated by the boxplot in Figure 4. This gives us 44 observations with around 11% censoring rate. We perform a goodness-of-fit test for the before-mentioned Power Generalized Weibull distribution $F_0^a(x, \sigma, v, \gamma) = F(x, \sigma, v, \gamma)$. Additionally, we also consider the Weibull distribution $F_0^b(x, \sigma, v) = F(x, \sigma, v, 1)$ and the Exponential distribution $F_0^c(x, \sigma) = F(x, \sigma, 1, 1)$, where

$$F(x, \sigma, v, \gamma) = 1 - \exp \left(1 - (1 + (x/\sigma)^v)^{1/\gamma} \right).$$

First, we fitted the parameters using MLE under random censoring obtaining the estimators $(\hat{\sigma}, \hat{v}, \hat{\gamma})$ and the following distributions $F_0^a(x, 4.63, 1.82, 1.91)$, $F_0^b(x, 1.44, 8.45)$ and $F_0^c(x, 8.33)$. Then we applied our test J_n and obtained the following p-values: $p^a = 0.86$, $p^b = 0.88$ and $p^c = 0.01$ for Power Generalized Weibull, Weibull and Exponential, respectively. Hence, the results of the test confirm what it is to see in Figure 4, that both Power Generalized Weibull and Weibull fit the data very well while the Exponential distribution is not adequate to describe the head-and-neck cancer data.

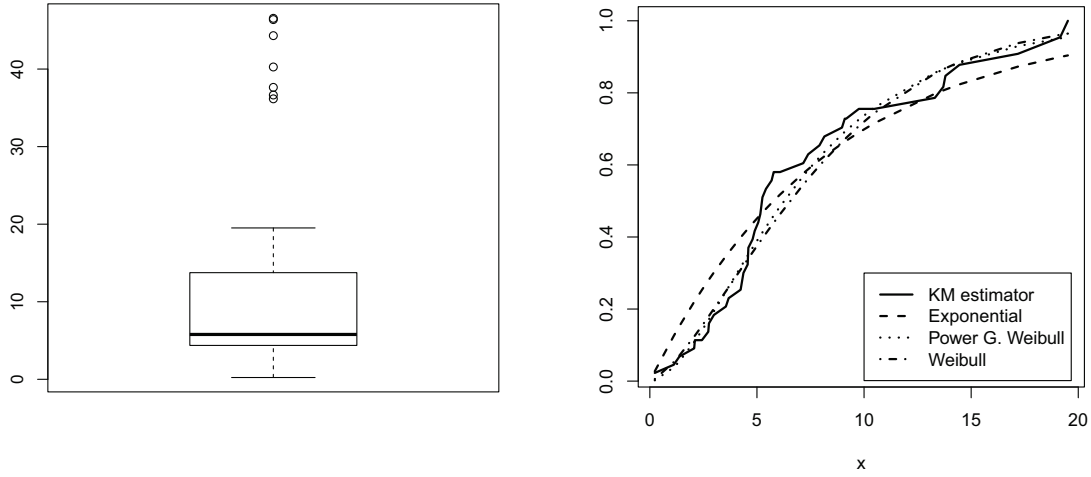


Figure 4: Boxplot (left) and Kaplan-Meier estimator (right) together with $F_0^a(x, \hat{\sigma}, \hat{v}, \hat{\gamma})$ (dashed), $F_0^b(x, \hat{\sigma}, \hat{v})$ (dotted) and $F_0^c(x, \hat{\sigma})$ (dot-dashed) for the head-and-neck cancer data.

7 Conclusions

In this work we developed and studied the goodness-of-fit test based on maximum correlation under random censoring. The advantage of our test over other goodness-of-fit tests, like χ^2 test proposed by Akritas (1988), is its simplicity. Our test bases on asymptotic normality and neither number of classes nor the class boundaries have to be chosen. The proposed test outperforms the one by Akritas (1988) for most of the alternatives studied. Even though the test was designed to check the uniformity, with simple transformation it can be applied for any, fully specified continuous distribution. Finally, it can be extended to composite hypothesis, that is, when the distribution in the null hypothesis is known up to a parameter. The jackknife modification to the asymptotic variance has been proposed. A theoretical study on the test under the composite null hypothesis is out of the scope of the present paper and purpose of further research.

Acknowledgements

This work has been partially supported by Spanish grants MTM2010-17323 (Spanish Ministry of Science and Innovation), MTM2011-22392, ECO2011-25706 (Spanish Ministry of Economy and Competitiveness).

References

- [1] Akritas, M. G. (1988) *Pearson-Type Goodness-of-Fit test: The Univariate Case*. Journal of the American Statistical Association 83, pp. 222-230

- [2] Balakrishnan, N., Chimitova E., Vedernikova, M. (2014) *An empirical analysis of some non-parametric goodness-of-fit tests for censored data* Communications in Statistics - Simulation and Computation, DOI: 10.1080/03610918.2013.796982
- [3] Cuadras, C. M., Fortiana, J.(1993) *Continuous metric scaling and prediction* Multivariate Analysis, Future Directions, vol. 2 (eds C. M. Cuadras and C. R. Rao), pp. 47-66. Amsterdam: North-Holland.
- [4] Fortiana, J., Grané, A. (2003) *Goodness-of-Fit Tests Based on Maximum Correlations and Their Orthogonal Decompositions* J. Royal Statistical Society. Series B 65, pp. 115-126.
- [5] Grané, A. (2012) *Exact goodness-of-fit tests for censored data* Ann. Inst. Stat. Math. 64, pp.1187-1203.
- [6] Haghighi, F., Nikulin, H. (2004) *A chi-squared test for power generalized Weibull family for the head-and-neck cancer censored data* Zapiski Naucznych Sjeminarov 311, pp. 222-236.
- [7] Kaplan, E. L., Meier, P. (1958) *Nonparametric estimation from incomplete observations* Journal of the American Statistical Association 53, pp. 457-481.
- [8] Koziol J.A., Green S.B. (1976) *A Cramer-von Mises Statistic for Randomly Censored Data* Biometrika 63, pp. 465-474.
- [9] Rao, R.R. (1962) *Relations Between Weak and Uniform Convergence of Measures with Applications* Annals of Mathematical Statistics 33, pp. 659-680.
- [10] Stute, W.(1994) *The Bias of Kaplan-Meier Integrals* Scandinavian Journal of Statistics 21, pp. 475-484.
- [11] Stute, W.(1995) *The Central Limit Theorem Under Random Censorship* Annals of Statistics 23, pp. 422-439.
- [12] Wellner, J.A. (2007) *On an exponential bound for the Kaplan-Meier estimator* Lifetime Data Analysis 13, pp.481-496.
- [13] Ying, Z. (1989) *A note on the asymptotic properties of the product-limit estimator on the whole line* Statistics & Probability Letters 7, pp. 311-314.